Protein Engineering vol.3 no.3 pp.173–180, 1990

Crystallization and structure solution at 4 Å resolution of the recombinant synthase domain of *N*-(5′-phosphoribosyl)anthranilate isomerase:indole-3-glycerol-phosphate synthase from *Escherichia coli* complexed to a substrate analogue

Matthias Wilmanns, Edith Schlagenhauf, Bruno Fol and Johan N.Jansonius[1]

Department of Structural Biology, Biocentre, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

[1]To whom correspondence should be addressed

The recombinant synthase domain of the bifunctional enzyme *N*-(5′-phosphoribosyl)anthranilate isomerase:indole-3-glycerol-phosphate synthase from *Escherichia coli* has been crystallized, and the structure has been solved at 4 Å resolution. Two closely related crystal forms grown from ammonium sulphate diffract to 2 Å resolution. One form (space group *R*32, a = 163 Å, α = 29.5°) contains the unliganded synthase domain; the second crystal form (space group $P6_322$, a = 144 Å, c = 158 Å) is co-crystallized with the substrate analogue *N*-(5′-phosphoribit-1-yl)anthranilate. The structure of the synthase–inhibitor complex has been solved by the molecular replacement method. This achievement represents the first successful use of a $(\beta\alpha)_8$-barrel monomer as a trial model. The recombinant synthase domain associates as a trimer in the crystal, the molecules being related by a pseudo-crystallographic triad. The interface contacts between the three domains are mediated by those residues that are also involved in the domain interface of the bifunctional enzyme. This system provides a model for an interface which is used in both intermolecular and intramolecular domain contacts.

Key words: $(\beta\alpha)_8$-barrel/crystal structure/IGP synthase/interface contacts/molecular replacement

Introduction

The monomeric bifunctional enzyme *N*-(5′-phosphoribosyl)-anthranilate isomerase:indole-3-glycerol-phosphate synthase (PRAI:IGPS) from *Escherichia coli* (EC 4.1.1.48) catalyses two sequential reactions in the biosynthesis of tryptophan. *N*-(5′-phosphoribosyl)anthranilate (PRA) is converted by PRAI to 1-(*o*-carboxyphenylamino)-1-deoxyribulose 5-phosphate (CdRP) via a practically irreversible Amadori rearrangement (Kirschner *et al.*, 1987). The succeeding step, the ring-closure reaction from CdRP to indole-3-glycerol-phosphate (IGP), is catalysed by IGPS. It has been shown by biochemical and genetic studies that the N-terminal half of this bifunctional enzyme is responsible for the IGPS activity, whereas the C-terminal half catalyses the PRAI reaction (Yanofsky *et al.*, 1971; Cohn *et al.*, 1979; Kirschner *et al.*, 1980). The amino acid sequence inferred from the gene sequence consists of 452 amino acid residues (Horowitz *et al.*, 1983).

The crystal structure of PRAI:IGPS was solved at 2.8 Å resolution by the multiple isomorphous replacement method (Priestle *et al.*, 1987). Using oscillation data collected at DESY, Hamburg, the model has been partially refined to a current *R*-factor of 27% at 2.4 Å resolution (M.Wilmanns, unpublished results). The enzyme contains two distinct functional domains, each folded as a $(\beta\alpha)_8$-barrel. The domain that catalyses the IGPS reaction has an additional N-terminal segment of 47 amino acid residues. The active sites have been approximately located using an iodinated substrate analogue and were found to be in depressions on the surface of the domains created by the outward-curving loops between C-termini of the β-strands and the subsequent α-helices (Priestle *et al.*, 1987).

The two domains of PRAI:IGPS from *E.coli* were cloned and expressed separately by site-directed mutagenesis by Pflugfelder (1986). TGT, which codes for Cys260, was replaced by the opal-type stop codon TGA into *trpC(F)* at that position. The IGPS domain ($IGPS_{mon'}$) was expressed in *E.coli* by this mutated *trpC(F)* gene on a plasmid. It has been shown by steady-state kinetics that $IGPS_{mon'}$ has synthase activity of the same order of magnitude as in the native bifunctional enzyme (M.Eberhard, personal communication).

Here we describe two crystallographically closely related crystal forms of $IGPS_{mon'}$. For one of these, a complex between $IGPS_{mon'}$ and the substrate analogue *N*-(5′-phosphoribit-1-yl)anthranilate (rCdRP), the crystal structure has been determined by molecular replacement methods. Because rCdRP differs from CdRP only by selective reduction of the 2′-carbonyl group to an alcohol group, it can be regarded as a close substrate analogue (Bisswanger *et al.*, 1979). It is a competitive inhibitor for IGPS with K_i = 8 μM in 0.1 M potassium phosphate buffer, pH 7.5, at 20°C.

Materials and methods

Purification

$IGPS_{mon'}$ was isolated from a homogenate of *E.coli* cells that had been transformed with a recombinant plasmid containing the gene for this domain (Pflugfelder, 1986). This gene *trpC′* codes for residues 1–259 of the native bifunctional enzyme. The protein purification was accomplished by an existing procedure (Pflugfelder, 1986). The last purification step (gel permeation by Sephacryl S-200) was omitted as the enzyme was at least 95% pure following the preceding steps. The purity of each preparation was checked by PAGE under denaturing conditions (Laemmli, 1970). The specific activity was determined according to Kirschner *et al.* (1987).

Crystallization

Purified protein was dialysed against 0.05 M potassium phosphate (pH 7.5, 4°C) containing 0.005 M EDTA, 0.002 M 1,4-dithioerythritol and 0.02% (w/v) NaN_3 and was concentrated by centrifugation through a semipermeable membrane (Centricon, TM 10) to 10–15 mg/ml protein. Drops were prepared at 8°C by mixing equal volumes (5 μl) of freshly concentrated protein solution and reservoir buffer solution containing 0.05 M potassium phosphate (pH 5.0), 1.2 M ammonium sulphate, 0.005 M EDTA and 0.002 M 1,4-dithioerythritol. rCdRP (2 μl 0.1 M) was added to the drop for the crystallization of inhibited $IGPS_{mon'}$. The rCdRP was synthesized according to Bisswanger *et al.* (1979).

Two related crystal forms grew in ~4 weeks by the hanging drop method. The size of the crystal form II containing rCdRP liganded to $IGPS_{mon'}$ was increased by the macroseeding technique (Thaller *et al.*, 1981). The crystals were washed in 0.05 M potassium phosphate buffer (pH 6.0) for 2 h and then transferred to drops with the same buffer composition but with a reduced protein concentration of 5–10 mg/ml. The diffraction properties and space group assignments of both crystal forms were determined from precession photographs.

Data collection and processing

Oscillation data of the crystal form II were collected on films (Kodak, NS-5T) using an Arndt–Wonacott rotation camera (Enraf-Nonius) and graphite monochromated Cu-K_α radiation from a rotating anode generator operating at 40 kV and 40 mA. The crystals were mounted in thin-walled glass capillaries with the *a* axis oriented parallel to the spindle axis, requiring 90° of data to be collected. More efficient data acquisition would have resulted if the 6-fold *c* axis had been aligned along the spindle, but this was impractical due to the thinness of the crystals.

The X-ray beam was collimated such that the entire crystals were irradiated. The film packs (two films/pack) were placed at a distance of 140 mm from the crystal allowing data collection to 4 Å resolution. The crystals were cooled to 5°C in order to prolong their lifetime in the X-ray beam. An oscillation range of 2° per pack was used which required an exposure time of 8–11 h. Forty-eight film packs with partially overlapping data from six crystals completed a total range of 90°.

The data from the film packs were digitized on a rotating drum scanning densitometer (Optronics Photoscan System, P 1000) using a raster size of 100 μm and a 3.0 optical density scale. The data were processed using a version of the Cambridge oscillation program system (Nyborg and Wonacott, 1977) which incorporates the profile analysis algorithm of Rossmann (1979) as modified by Wilson *et al.* (1983). The data were scaled using the algorithm of Fox and Holmes (1966) and were refined using the program POSTREF (Winkler *et al.*, 1979) as modified by Greenhough and Helliwell (1982). All negative intensities were set to zero. The final data set included 8536 unique reflections constituting 92.5% of the unique data to 4 Å resolution. The final R_{merge} value for all data (where $R_{merge} = \Sigma_i \mid <I_i> - I_i \mid / \Sigma_i I_i$ and $<I_i>$ is the average of I_i over all symmetry equivalents) was 8%.

Structure determination

The structure was solved by the molecular replacement method taking the synthase domain of the partially refined bifunctional enzyme PRAI:IGPS (Priestle *et al.*, 1987; M.Wilmanns, unpublished results) as a trial model. The rotational search was carried out using the program ALMN (written by E.Dodson, York University) which is a version of the fast rotation function of Crowther (1972). A model of the synthase domain ($IGPS_{bi}$) was positioned in a *P*1 cell with arbitrarily chosen cell axes of 100 Å and cell angles of 90°. Structure factors were calculated to 4 Å resolution, assuming an overall temperature factor, *B*, of 25 Å^2. The Patterson search was carried out using a step size of 2.5° (in the Eulerian angle system α, β, γ) for the α and γ rotation angles. The step size for the β rotation was initially 5° and finally reduced to 0.5° in the vicinity of the solution. The Patterson space chosen was a shell of inner and outer radii of 5 and 30 Å. The search was carried out using data from various resolution shells between 4 and 10 Å.

The correctly oriented molecule was submitted to a translation search using the T_2 translation function of Crowther and Blow (1967), implemented in the computer program TFSGEN by Dr I.Tickle (Birkbeck College, London). The observed data between 10 and 4 Å resolution were included. In order to obtain an interpretable solution in the translation function a 'sub'-cell was used by selecting reflections with indices ($k = h \pm 3n$), which will be described in detail below (see Results).

Table I. Properties of crystals of IGP synthase domain

	Crystal form I	Crystal form II
rCdRP	absent	present
Space group	*R*32	$P6_322$
No. of molecules/ asymmetric unit	1	3
Cell dimensions		
Rhombohedral	a = 163 Å	
	α = 29.5°	
Hexagonal	a = 83 Å	a = 144 Å
	c = 465 Å	c = 158 Å
Cell volume		
Rhombohedral	924 000 Å^3	
Hexagonal	2 770 000 Å^3	2 840 000 Å^3

A trimer of IGPS molecules was generated from the model which arose from the solution of the translation function. It was refined by the least-squares program CORELS (Sussman *et al.*, 1977) in which each monomeric unit was treated as an independent rigid body. One cycle of scale factor refinement, three cycles including observed data between 10 and 8 Å resolution and one final cycle including data between 10 and 6 Å resolution were carried out.

Structure improvement

A Sim weighted ($2F_o - F_c$) electron density map (Sim, 1959) was computed using calculated phases and structure factor amplitudes derived from the CORELS refined model. In order to reduce the bias towards this model, maps were limited to small portions of the molecule (typically 5–8%), for which the contributions to the calculated phases and structure factor amplitudes were omitted. A set of 'omit'-maps covering a complete molecule was assembled by an automated procedure, implemented by Dr J.P.Priestle, Basel. In addition, a σ_a weighted electron density map (Read, 1986) was computed.

All three $IGPS_{mon'}$ domains were rebuilt independently on an interactive graphic display system (Evans and Sutherland PS 330) using the FRODO software package (Jones, 1978), further developed at Rice University, Tex. (Pflugrath *et al.*, 1984) and modified by Dr P.Evans (MRC, Cambridge, UK). The atomic positions of the corrected model were refined by the restrained least-squares refinement program package TNT (Tronrud *et al.*, 1987). In order to maintain a favourable ratio between observed data (7689 reflections between 8 and 4 Å resolution) and parameters to be refined (2038 non-hydrogen atoms per monomer) only the coordinates of one $IGPS_{mon'}$ domain were refined at a time while the coordinates of the other two monomers were fixed. This procedure was carried out twice for the trimer.

Results

Crystal characterization

Two closely related crystal forms of the recombinant $IGPS_{mon'}$ domain (residues 1–259) from *E.coli* grow from ammonium sulphate at pH 5.0 (Table I). By reindexing the rhombohedral crystal form I of the unliganded protein to a hexagonal crystal

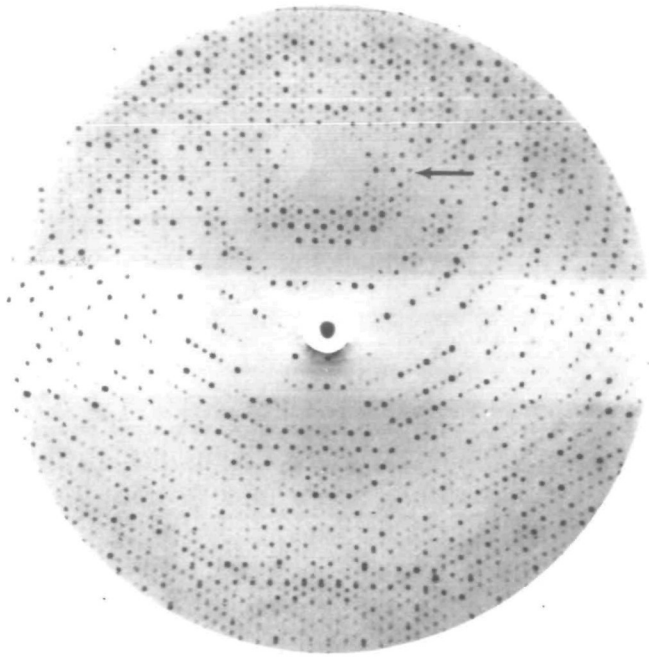

Fig. 1. Oscillation photograph covering the rotation range $\phi = 12° - 14°$ ($\phi_{hk0} = 0°$, $\phi_{h0l} = 90°$). Experimental details are given in Materials and methods. The systematic difference between the intensities of the strong reflections ($h = k \pm 3n$) and the remaining reflections ($h \neq k \pm 3n$) can be seen, for example, in the line of the reflections which are indexed *hk*2 (marked by an arrow).

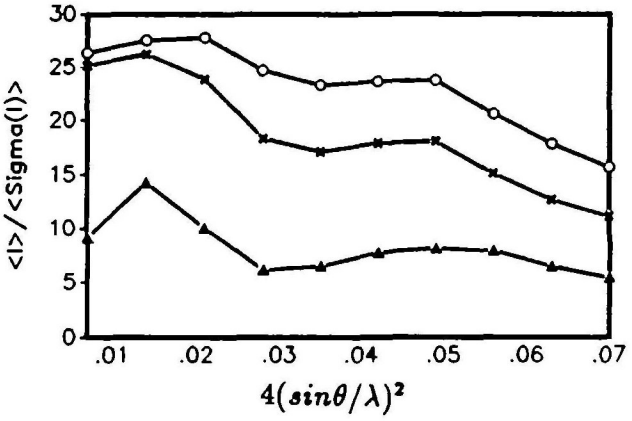

Fig. 2. $<I>/<\sigma(I)>$ as a function of resolution for all data to 4 Å resolution. ○, data $h = k \pm 3n$; ▲, data $h \neq k \pm 3n$; ×, all data.

system it appears that its *c* axis (hexagonal) is approximately three times longer than the *c* axis of the crystal form II which contains the liganded protein. The *ab* plane (hexagonal) of I is about one-third the area of the *ab* plane of II. The difference between the cell volumes of I and II is <3%. Both crystal forms diffract to ~2.0 Å resolution and are therefore suitable for X-ray analysis.

Crystal form I grows as hexagonal prisms with a maximum size of 0.6 × 0.6 × 0.3 mm^3. Crystal form II, which contains rCdRP liganded to $IGPS_{mon'}$, grows with maximum dimensions of 0.5 × 0.5 × 0.1 mm^3. These crystals have the shape of an optical lens.

Absorption spectra of crystal form II were recorded with a microspectrophotometer (Zeiss, 03) controlled by a micro-computer (Hewlett-Packard, 9845B) to ascertain whether rCdRP was specifically bound to the crystallized protein. Free rCdRP in solution has an absorption maximum at 327 nm which is shifted

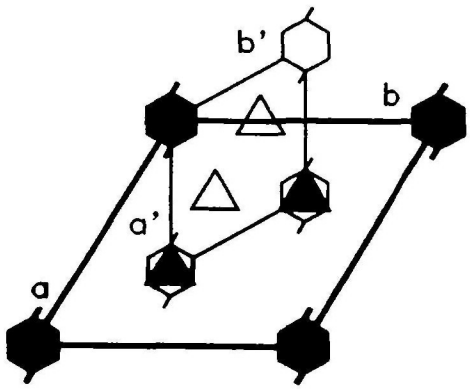

Fig. 3. Relation between the large cell (thick lines, filled symmetry symbols) and the subcell (thin lines, open symmetry symbols). The axes of the large cell are labelled (*a*,*b*) and of the subcell (*a'*,*b'*) respectively. The subcell is rotated by 30° around *c*, relative to the large cell. All additional 2-fold axes and translation components have been omitted for clarity. The local triads of the large cell are located at the positions of the crystallographic 3-fold axes in the subcell.

by ~40 nm to longer wavelengths upon binding to PRAI:IGPS (Bisswanger *et al.*, 1979, Figure 8A). The absorption spectrum of the crystal form II shows a shift of the peak maximum of ~60 nm, providing evidence that rCdRP is specifically liganded to $IGPS_{mon'}$. Crystal form I does not absorb significantly at wavelengths longer than 320 nm.

Analysis of the diffraction pattern

A relation between the two crystal forms is evident from the diffraction pattern of crystal form II. It shows a pattern of strong intensities for reflections with indices ($k = h \pm 3n$) which corresponds to a 'sub'-cell of the same space group with axis $a' = 83$ Å $= a/\sqrt{3}$. All other reflections ($h \neq k \pm 3n$) are considerably weaker (Figure 1). $<I>/<\sigma>$ is 22.9 for the 2988 reflections ($h = k \pm 3n$) to 4 Å resolution, whereas this value drops to 7.5 for the remaining 5500 reflections ($<I>$ is the mean intensity and $<\sigma>$ is the mean standard deviation on intensities). The analysis of $<I>/<\sigma>$ as a function of resolution is given in Figure 2. While only 2.6% of the reflections ($h = k \pm 3n$) have negative intensities, this fraction increases to 12.8% for the remaining data ($h \neq k \pm 3n$). No such anomalies were observed in the intensity distribution as function of the index *l*.

In crystal form I the reflections of every third layer of index *l* (hexagonal system) are equivalent to reflections which correspond to the subcell of crystal form II. The presence of reflections in the remaining *l* layers (hexagonal system) requires the assignment of crystal form I to a rhombohedral crystal lattice (Table I).

Packing arrangement

A packing arrangement of the $IGPS_{mon'}$ domains in crystal form II was considered by taking into account the anomalous intensity distribution of the diffraction pattern. There are three molecules in the asymmetric unit of the real cell, which will be called 'large' cell in the remaining text for ease of discussion. The molecules are related by a local triad which deviates only slightly from crystallographic symmetry. This triad might be distorted by breaking the exact C_3 symmetry and/or by a tilt with respect to the crystallographic 3-fold rotation and 6-fold screw axis. In other words, there is a trimer of $IGPS_{mon'}$ domains per asymmetric unit related by pseudo-crystallographic symmetry. The triad is located at ($x = 1/3$, $y = 1/3$), which is equivalent to the position of a crystallographic 3-fold axis in crystal form I, since the

cell of II is rotated by 30° relative to the cell of I (Figure 3). Additional local triads are generated by application of symmetry operators.

The local 3-fold symmetry of II was replaced by crystallographic symmetry by setting the intensities of all reflections with indices ($h \neq k \pm 3n$) equal to zero. Omission of these data corresponds essentially to use of the subcell, neglecting the fact that the intensities of the strong reflections ($h = k \pm 3n$) also carry information on deviations from ideal symmetry. The subcell belongs to the same space group $P6_322$ as the large cell, but does not contain additional non-crystallographic symmetry elements. It is one-third of the volume of the large cell and contains one molecule instead of three in the asymmetric unit. The *ab* plane of the subcell of II is essentially equivalent to the *ab* plane of I, highlighting the close relations between both crystal forms (Table I).

These arguments were corroborated by an estimate of the packing density by the method of Matthews (1968). The mol. wt of $IGPS_{mon'}$ as calculated from the amino acid sequence is M_r = 28 933 (Pflugfelder, 1986). Three molecules per asymmetric unit result in a V_M = 2.7 Å^3/dalton, corresponding to a solvent content of 54% which is in the range commonly found for proteins.

Structure solution

The structure of liganded $IGPS_{mon'}$ (crystal form II) was solved by the molecular replacement method using the synthase domain of the bifunctional enzyme (PRAI:IGPS) as search model. Pure $(\beta\alpha)_8$-barrels have approximate 8-fold symmetry (Lebioda *et al.*, 1982), which is reduced to 4-fold symmetry when taking into account the side chain orientations of the residues which belong to the central β-barrel (Lesk *et al.*, 1989). Due to this symmetrical appearance of $(\beta\alpha)_8$-barrels it was expected that additional peaks could arise in the rotation function and impede the interpretation. Despite the large number of $(\beta\alpha)_8$-barrel structures available (Chothia, 1988), to our knowledge no crystal structure consisting of a monomeric $(\beta\alpha)_8$-barrel has been determined by the molecular replacement method. Here the trial model, $IGPS_{bi}$, is a single $(\beta\alpha)_8$-barrel. However, it contains a segment of 47 residues (18% of all residues) at the N-terminus in addition to the $(\beta\alpha)_8$-barrel which diminishes the symmetrical appearance of this structure.

It is possible under most conditions to reveal the nature of local symmetry by using the self-rotation function. However, the failure of the self-rotational search (program ALMN) was not unexpected due to the predicted pseudo-crystallographic arrangement of the local triads. Peaks corresponding to $IGPS_{mon'}$ domains related by these triads would be disguised by peaks resulting from the crystallographic threefold axes.

The cross-rotation function (program ALMN) gave single peaks of comparable strength for the subcell and the large cell, at positions related by rotations of 30° about the *c* axis. Three independent peaks would have been anticipated for a trimer in a general orientation, each peak describing the orientation of one monomer relative to the trial model. The appearance of single peak solutions supports the presumed pseudo-crystallographic nature of the local triad.

The peak heights are 5.0 and 4.9 σ above the mean for the large cell and the subcell respectively. Since the rotation function is dominated by diffraction data with strong intensities, it becomes clear that, in this case, the systematically weak reflections with indices ($h \neq k \pm 3n$) are *de facto* neglected. Hence the peaks for both cells arise essentially from the same data.

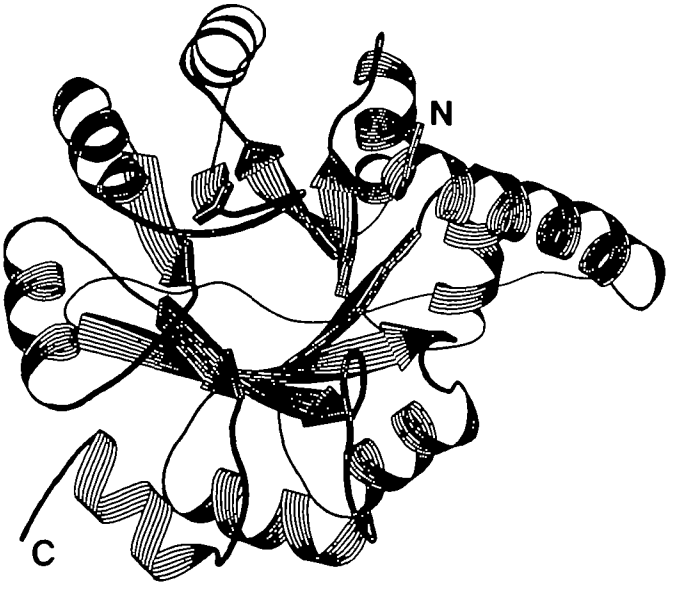

Fig. 4. Schematic drawing of the overall structure of $IGPS_{mon'}$, created with the program RIBBON (Priestle, 1988). The view is parallel to the axis of the central β-barrel. Helical ribbons, arrows and ropes represent α-helices, β-strands and coils respectively. The N-terminus starts in the upper right of the figure with a five-turn α-helix which covers the $(\beta\alpha)_8$-barrel horizontally. The $(\beta\alpha)_8$-barrel begins on the left side of the figure with the first β-strand (β_1) which is followed by the remaining seven strands in an anticlockwise rotation. Adjacent strands are connected by intervening α-helices shielding the central β-barrel from solvent. The C-terminus is located at the lower left of the figure. The nomenclature of secondary structure elements used in this publication has been taken from Priestle *et al.* (1987). The assignment of secondary structure elements to the sequence has been obtained from the partially refined model of $IGPS_{bi}$ (M.Wilmanns, unpublished results) and is not identical with the assignment in Figure 2 of Priestle *et al.* (1987). The present alignment is as follows: β_1, 47–55; α_1, 70–80; β_2, 83–88; α_2, 97–106; β_3, 110–116; α_3, 120–128; β_4, 133–137; α_4, 146–155; β_5, 158–164; α_5, 167–176; β_6, 180–184; α_6, 196–205; β_7, 209–214; α_7, 220–229; β_8, 233–236; α_8, 245–253.

The best signal-to-noise ratio was obtained by using reflections between 4 and 6 Å, the highest resolution available. It has been reported that in cases where pseudo-crystallographic symmetry or special packing arrangements are problematic, the best results were received by using higher resolution data (Dodson, 1985; Driessen and White, 1985).

The translation function (program TFSGEN) was initially applied using the subcell. This essentially replaces the assumed local triad by crystallographic symmetry, thereby simplifying the translational search such that only one peak was expected instead of three. In addition, it was thought that the signal-to-noise ratio would be worse using the large cell, since the search model would then constitute only one-third the scattering matter in the asymmetric unit. One peak (9.8 σ above the mean) gave the correct translation in the subcell and was applied to the coordinates of $IGPS_{bi}$. A trimer was generated by rotation about the presumed local triad at (x = 0, y = 1/3) (cf. Figure 3). A subsequent translational search with this trimer as trial model using the large cell gave the same solution, this time with peak height of 33.8 σ above the mean.

It should be noted that it is doubtful whether or not it would have been possible to solve the translation function with one search molecule per asymmetric unit in the large cell without the assumptions concerning the pseudo-crystallographic arrangement of the trimer. We are aware of two examples where knowledge of local symmetry was used for structure

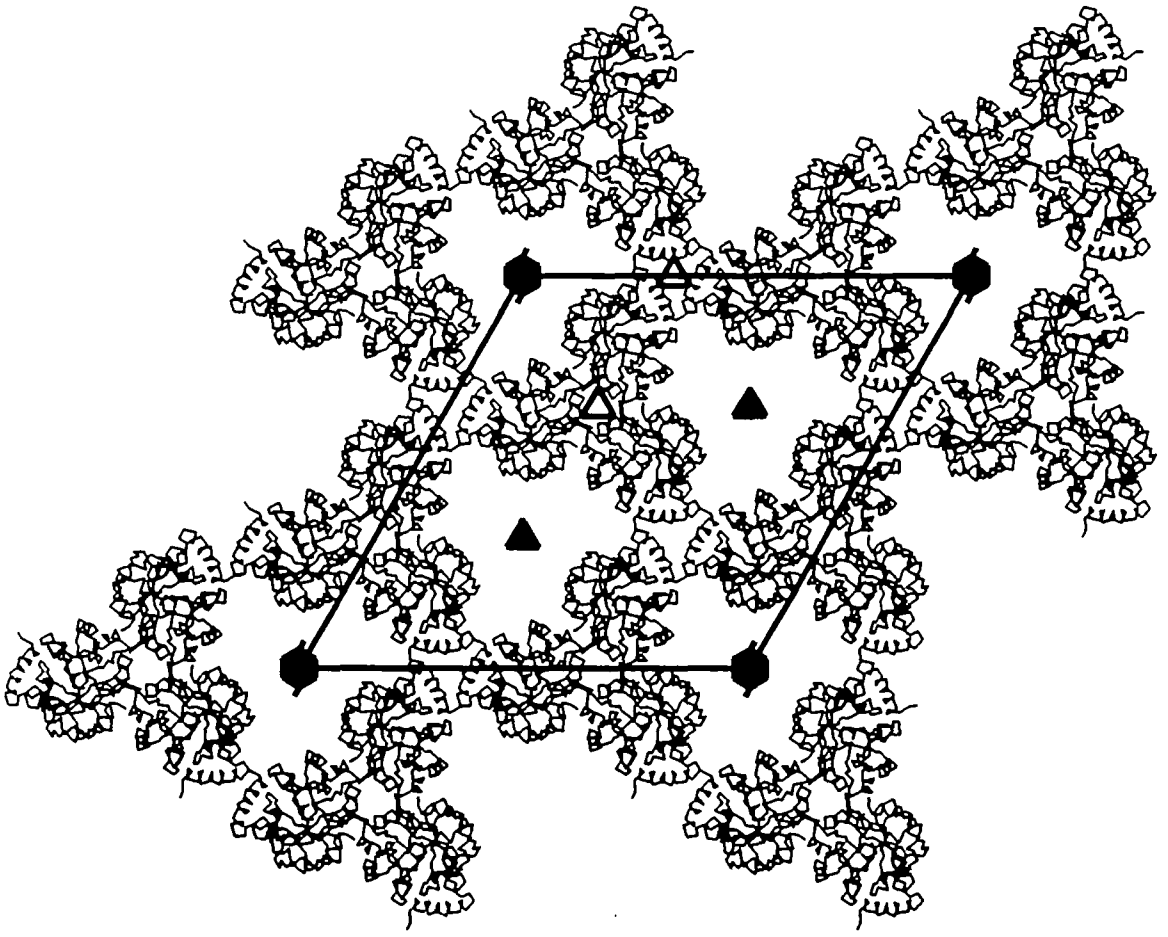

Fig. 5. The packing of $IGPS_{mon}$ domains in the crystal form II looking down the crystallographic *c* axis (same view as in Figure 3). Each $IGPS_{mon}$ domain is represented by its C_α backbone. All trimers of one layer in the *ab* plane are included, generated by the symmetry operations (x,y,x), $(-y,x-y,z)$, $(y-x,-x,z)$ and cell translations *a* and *b*. The crystallographic 3-fold rotation and 6-fold screw axes are labeled with filled symbols. Both types of local triads at positions (x = 0, y = 1/3) and (x = 1/3, y = 1/3) are labelled with open symbols (cf. Figure 3). The molecules above and below this layer are rotated by 60° about *c* relative to the molecules shown. The trimer interface is located near (x = 1/3, y = 1/3).

determination by molecular replacement. One case was reported by Schirmer *et al.* (1987) who solved the structure of hexameric C-phycocyanin from *Agmenellum quadruplicatum*. They deduced the presence of three $\alpha\beta$ units per asymmetric unit from the crystal properties and were able to locate the presumed trimer in the cell (space group *P*321). The structure of *γIVa*-crystallin from bovine lens was solved with the molecular replacement method by aligning a local twofold axis parallel to *c* (White *et al.*, 1988).

The rigid body refinement program CORELS moved the centres of the three $IGPS_{mon}$ domains over distances of up to 0.5 Å in *x*, 0.4 Å in *y* and 0.9 Å in *z*. Rotations around the centres of the molecules were <3°. The local triad was preserved although it was not restrained during refinement. The CORELS refined model gave an *R*-factor of 48% using all diffraction data between 10 and 4 Å resolution.

$(2F_o-F_c)$ electron density maps allowed unambiguous positioning of all main chain atoms (except the C-terminal residues 257–259) and 86% of the side chain atoms, even though the current data set is limited to 4 Å resolution. From the present model (after refinement by the least-squares program TNT) an *R*-factor of 28% was calculated for all data between 10 and 4 Å resolution.

Description of the crystal structure

The inhibited $IGPS_{mon}$ domain is folded as in the bifunctional enzyme (Priestle *et al.*, 1987). It contains a central $(\beta\alpha)_8$-barrel which was first found for triose phosphate isomerase (Banner *et al.*, 1975) and is often referred to as the 'TIM-barrel' (Figure 4). The barrel contains 207 residues starting with β_1 at Thr48. The preceding 47 residues wrap halfway around the compact $(\beta\alpha)_8$-barrel. The N-terminus is located at the C-terminal side of the central β-barrel and is followed by a five-turn helix (residues 4–22) which is closely packed against helix α_3 of the barrel. A long loop connects the end of this helix with the start of the $(\beta\alpha)_8$-barrel (Thr48).

The packing arrangement of the $IGPS_{mon}$ domains in the *ab* plane is shown in Figure 5. The trimer of $IGPS_{mon}$ domains in this crystal form is generated by a local triad parallel to *c* at (x = 1/3, y = 1/3). Several residues of the $(\beta\alpha)_8$-barrel are involved in the intermolecular contacts, yielding a tight interface within the trimer (Table II).

A second type of local triad at (x = 0, y = 1/3) is generated by applying crystallographic symmetry operations on the crystal trimer (Figure 5). The contacts around this local 3-fold axis are formed by residues from the long N-terminal helix. Equivalent residue ranges (17–21) belonging to different trimers of symmetry-related molecules face each other. Unfortunately, the electron density in this contact area is poor, leading to the assumption that this part of the crystal structure of $IGPS_{mon}$ is slightly disordered. These interactions between symmetry-related trimers are weaker and less numerous than those within the trimer interface.

Wide solvent channels with diameters of ~40 Å run parallel

to the *c* axis and are centred at the positions of the crystallographic 3-fold rotation and 6-fold screw axes (Figure 5).

There are clusters of intermolecular contacts where the local 3-fold symmetry breaks down. A complex pattern of interactions is established by identical side chains (Arg224, Ser227, His228) within the trimer interface (Table II). It can be assumed that the guanidinium groups of Arg224 of three $IGPS_{mon'}$ domains, which are close to the local triad, would interact via hydrogen bonds in a cyclic way. However, such a cluster of contacts would result in an unfavourable accumulation of three positive charges. These interactions will be examined in greater detail when the structure of $IGPS_{mon'}$ is refined at high resolution.

A (F_o-F_c) map shows strong electron density for the substrate analogue rCdRP at the presumed active site (Priestle *et al.*, 1987), between residues Glu53, Glu214, Ser215 and Ser237. These residues are partially conserved in 11 sequences of IGPS (Priestle *et al.*, 1987). However, due to the limited resolution of the current data, an unambiguous fit of the rCdRP molecule to this electron density has not yet been possible.

Table II. Intermolecular contacts between $IGPS_{mon'}$ domains

	Mol. *i*	Mol. *i*+1	Class of contact	Hydrogen bonds	Breakage of local symmetry[a]
1[b]	Arg47	Leu205			
		Gly206	T		
2	His79	Glu166	T	+	
3	Tyr80	Glu200	T	+	
4	Arg199	Leu254	T		
5	Tyr220	Phe229	T		
6	Arg224	Arg224	T	+	*
7	Arg224	His228	T	+	*
8	Ser227	His228	T	+	*
9	Gln2	Lys91	C	+	
10[b]	Glu17	Glu17			
	Lys20	Lys20			
	Gln21	Gln21	C		
11	Val63	His242	C	+	
12	Asp66	Arg251	C	+	
13	Asp66	Asp243	C	+	
14	Ala241	Ala241	C		

Contacts between $IGPS_{mon'}$ domains within the crystal trimer are labelled T; contacts between crystallographic symmetry-related trimers are labelled C.
[a]If the pattern of contacts between molecule *i* and molecule *i*+1 is different for the three monomers, the contact is labelled *.
[b]The electron density in this contact area is poor, hence a detailed interpretation has not yet been possible.

Discussion

Structure comparison

In order to discuss global structural differences between $IGPS_{mon'}$ and $IGPS_{bi}$ the overall errors of the atomic positions were estimated by plotting *R*-factors against resolution (Luzatti, 1952). This resulted in an error estimate of 0.5 Å in the atomic positions of the current model.

The r.m.s. deviations of all 256 C_α atoms of the current model in the three pairs of superimposed $IGPS_{mon'}$ domains are 0.63, 0.75 and 0.79 Å. If the superpositions are restricted to the C_α atoms which belong to the β-strands of the barrel (34 atoms), the r.m.s. deviations drop to 0.29, 0.30 and 0.34 Å. The local triad is exactly parallel to *c*, and adjacent molecules in the trimer are related by rotations of 120 ± 2° about *c*. There is a small but significant translation component between each pair of $IGPS_{mon'}$ domains along the local triad. The values of the translations are +0.4, +0.4 and −0.8 Å, with an error of ~0.1 Å. This is one manifestation of the broken C_3 symmetry of the local triad in crystal form II.

Another set of comparisons was carried out by superimposing each individual $IGPS_{mon'}$ molecule of the trimer onto the $IGPS_{bi}$

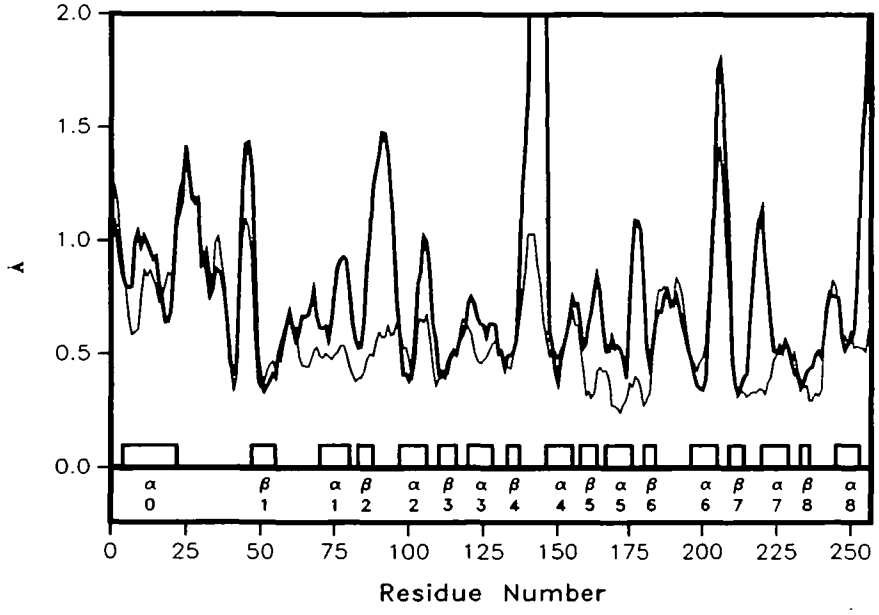

Fig. 6. Deviations in the C_α atom positions between pairs of superimposed $IGPS_{mon'}$ and $IGPS_{bi}$ molecules in comparison to the deviations in the C_α atom positions between superimposed pairs of the three $IGPS_{mon'}$ monomers. Heavy line: averaged differences between C_α atoms ($IGPS_{mon'}/IGPS_{bi}$). The values of residues 141−145 exceed the area of the graph; the maximum difference is 4.1 Å (Asp142). Thin line: averaged differences between C_α atoms of the three pairs ($IGPS_{mon'}/IGPS_{mon'}$). The curves have been smoothed by computing the value per residue (*i*) and averaging it over the neighbouring five residues [(*i* − 2) . . . (*i* + 2)]. The assignments of the secondary structure elements are given at the bottom of the figure (cf. Figure 4). The N-terminal helix is labelled α_0. Loops are named in the text according to the secondary structure elements which connect them, e.g. $\beta_1\alpha_1$. The last three residues at the C-terminus (residues 257−259) are not included in the current model.

domain from the bifunctional enzyme. The overall r.m.s. values for all C_α atoms were 1.08, 1.10 and 1.17 Å. There are some segments where the average differences between C_α atoms ($IGPS_{mon'}/IGPS_{bi}$) are considerably larger than those between C_α atoms ($IGPS_{mon'}/IGPS_{mon'}$), especially in the barrel loops $\beta_2\alpha_2$ and $\beta_4\alpha_4$ (Figure 6). Since both loops contribute to the active site of IGPS and contain several invariant residues (Priestle *et al.*, 1987), these deviations may arise from binding of the substrate analogue rCdRP to the active site of IGPS. Further significant differences were observed for the carboxyl end of strand β_5 (Glu163 is invariant) as well as for the loops $\alpha_5\beta_6$, $\beta_7\alpha_7$ and helix α_7. For the N-terminal part of the structure, the carboxyl end of helix α_6 and loop $\alpha_6\beta_7$, the differences between C_α atoms are generally high in both types of superpositions, $IGPS_{mon'}/IGPS_{mon'}$ and $IGPS_{mon'}/IGPS_{bi}$. The electron density of a $(2F_o-F_c)$ map is generally weak in these areas so that these segments are considered to be rather flexible in the crystal structure of $IGPS_{mon'}$.

Intermolecular contacts

The interface between the IGPS domain and the PRAI domain of the partially refined model of the bifunctional enzyme was examined for intramolecular domain−domain contacts. The following IGPS residues are closer than 4 Å to the PRAI domain: 47−50, 220, 223−224, 247 and 250−254. The interface includes seven hydrogen bonds plus a number of hydrophobic contacts.

All of these residues except Ala247 are involved in intermolecular contacts within the trimer interface of $IGPS_{mon'}$ molecules (Figure 7). In contrast to this, of all the residues (92, 155, 157, 186, 193, 228) within $IGPS_{bi}$ that contribute to crystal contacts of the bifunctional enzyme, only one (His228) makes intermolecular contacts between $IGPS_{mon'}$ molecules. We conclude that the residues involved in contacts between the two domains of $IGPS_{bi}$ also form the crystal trimer interface of the $IGPS_{mon'}$ molecules. It can be regarded as a quasi-subunit interface. Solution studies by gel permeation chromatography

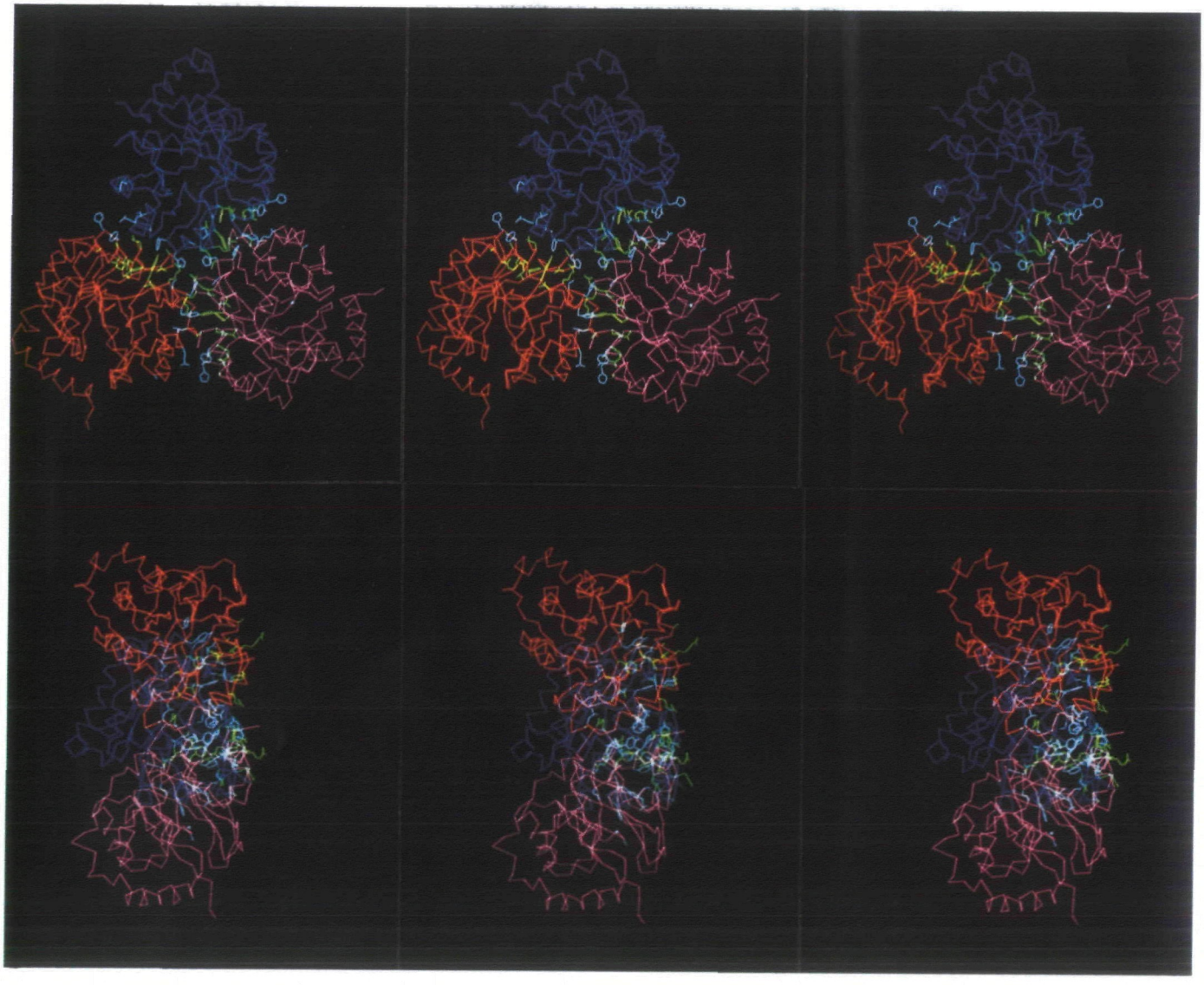

Fig. 7. Stereo presentation of the trimeric association of $IGPS_{mon'}$ highlighting the interface between the three monomers. The C_α backbones of $IGPS_{mon'}$ (1), $IGPS_{mon'}$ (2) and $IGPS_{mon'}$ (3) are in red, pink and blue respectively. Side chains of trimer residues which are involved in the interface of $IGPS_{bi}$ are shown in yellow. Side chains of residues which make further contacts in the trimer interface are shown in white. (**A**) Stereo image of the trimer viewed down the crystallographic *c* axis; (**B**) stereo image of the trimer with the *c* axis horizontally oriented. The contacts are concentrated on one side (at the right in this picture), resulting in an overall tripod-like shape of the trimer.

indicated an increased tendency for association of $IGPS_{mon'}$ compared to the bifunctional enzyme, but a definite trimeric association was not detected (M.Eberhard, personal communication).

These findings support the ideas of Argos (1988) and Miller *et al.* (1987), who have systematically studied the nature and composition of domain interfaces of monomeric enzymes and subunit interfaces of oligonmeric enzymes. These authors conclude that both types of interfaces have about the same amino acid compositions, which can be considered as lying between the two compositional extremes provided by interior and exterior amino acids. It is energetically unfavourable for large surfaces containing mainly hydrophobic residues to be exposed to an aqueous environment. In this paper we present an example where the domain interface contacts of a two-domain enzyme PRAI:IGPS can be directly compared to the intermolecular interface contacts of one of the two domains ($IGPS_{mon'}$), which is trimeric in the present crystal structure. The fact that almost all residues contributing to interdomain contacts in $IGPS_{bi}$ are also involved in the interface between $IGPS_{mon'}$ molecules is consistent with the statistical conclusions of Argos and Miller *et al.* It would be energetically favourable for the domain interface of $IGPS_{bi}$ to participate in $IGPS_{mon'}$ as a quasi-subunit interface rather than being exposed to solvent. A comparison of both structures at high resolution should allow a more quantitative analysis of these interfaces.

Acknowledgements

We thank Professor K.Kirschner for providing us with the *E.coli* strain KK8 and the plasmid containing the *trpC'* gene. We are grateful to M.Eberhard for advice on protein purification. Drs J.P.Priestle and R.A.Pauptit are thanked for helpful discussions concerning the structure solution. This work was supported by grants from the Swiss National Science Foundation (3.098-85 and 31-25712.88 to J.N.J.).

References

Argos,P. (1988) *Protein Engng*, **2**, 101–113.
Banner,D.W., Bloomer,A.C., Petsko,G.A., Phillips,D.C., Pogson,C.I., Wilson,I.A., Corran,P.H., Furth,A.J., Milman,J.D., Offord,R.E., Priddle,J.D. and Waley,S.G. (1975) *Nature*, **255**, 609–614.
Bisswanger,H., Kirschner,K., Cohn,W., Hager,V. and Hansson,E. (1979) *Biochemistry*, **18**, 5946–5953.
Chothia,C. (1988) *Nature*, **333**, 598–599.
Cohn,W., Kirschner,K. and Paul,C. (1979) *Biochemistry*, **18**, 5953–5959.
Crowther,R.A. (1972) In Rossmann,M.G. (ed.), *The Molecular Replacement Method*. Gordon and Breach, New York, pp. 173–178.
Crowther,R.A. and Blow,D.M. (1967) *Acta Crystallogr.*, **23**, 544–548.
Dodson,E.J. (1985) In Machin,P.A. (ed.), *Molecular Replacement*. Daresbury Laboratory, pp. 33–45.
Driessen,H.P.C. and White,H. (1985) In Machin,P.A. (ed.), *Molecular Replacement*. Daresbury Laboratory, pp. 27–32.
Fox,G.C. and Holmes,K.C. (1966) *Acta Crystallogr.*, **20**, 886–891.
Greenhough,T.J. and Helliwell,J.R. (1982) *J. Appl. Crystallogr.*, **15**, 493–508.
Horowitz,H., van Arsdell,J. and Platt,T. (1983) *J. Mol. Biol.*, **169**, 775–797.
Jones,T.A. (1978) *J. Appl. Crystallogr.*, **11**, 268–272.
Kirschner,K., Szadkowski,H., Henschen,A. and Lottspeich,F. (1980) *J. Mol. Biol.*, **143**, 395–409.
Kirschner,K., Szadkowski,H., Jardetzky,T.S. and Hager,V. (1987) *Methods Enzymol.*, **142**, 386–397.
Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
Lebioda,L., Hatada,M.H., Tulinsky,A. and Mavridis,I.M. (1982) *J. Mol. Biol.*, **162**, 445–458.
Lesk,A.M., Bränden,C.I. and Chothia,C. (1989) *Proteins*, **5**, 139–148.
Luzatti,V. (1952) *Acta Crystallogr.*, **5**, 802–810.
Matthews,B.W. (1968) *J. Mol. Biol.*, **33**, 491–497.
Miller,S., Janin,J., Lesk,A.M. and Chothia,C. (1987) *J. Mol. Biol.*, **196**, 641–656.
Nyborg,J. and Wonacott,A.J. (1977) In Arndt,U.W. and Wonacott,A.J. (eds), *The Rotation Method in Crystallography*. North-Holland, New York, pp. 139–152.
Pflugfelder,M. (1986) Dissertation, University of Basel, Switzerland.
Pflugrath,J.W , Saper,M.A. and Quiocho,F.A. (1984) in Hall,S. and Ashida,T. (eds), *The Methods and Applications in Crystallographic Computing*. Oxford University Press, London, p. 404.
Priestle,J.P. (1988) *J. Appl. Crystallogr.*, **21**, 572–576.
Priestle,J.P., Grütter,M.G., White,J.L., Vincent,M.G., Kania,M., Wilson,E., Jardetzky,T.S., Kirschner,K. and Jansonius,J.N. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 5690–5694.
Read,R.J. (1986) *Acta Crystallogr.*, **A42**, 140–149.
Rossmann,M.G. (1979) *J. Appl. Crystallogr.*, **12**, 225–238.
Schirmer,T., Huber,R., Schneider,M., Bode,W., Miller,M. and Hackert,M.L. (1986) *J. Mol. Biol.*, **188**, 651–676.
Sim,G.A. (1959) *Acta Crystallogr.*, **12**, 813–815.
Sussman,J.L., Holbrook,S.R., Church,G.M. and Kim,S.-H. (1977) *Acta Crystallogr.*, **A33**, 800–804.
Thaller,C., Weaver,L.H., Eichele,G., Wilson,E., Karlsson,R. and Jansonius,J.N. (1981) *J. Mol. Biol.*, **147**, 465–469.
Tronrud,D.E., Ten Eyck,L.F. and Matthews,B.W. (1987) *Acta Crystallogr.*, **A43**, 489–501.
White,H.E., Driessen,H.P.C., Slingsby,C., Moss,D.S., Turnell,W.G. and Lindley,P.F. (1988) *Acta Crystallogr.*, **B44**, 172–178.
Wilson,K.S., Stura,E.A., Wild,D.L., Todd,R.J., Stuart,D.I., Babu,Y.S., Jenkins,J.A., Standing,T.S., Johnson,L.N., Fourme,R., Kahn,R., Gadet,A., Bartels,K.S. and Bartunik,H.D. (1983) *J. Appl. Crystallogr.*, **16**, 28–41.
Winkler,F.K., Schutt,C.E. and Harrison,S.C. (1979) *Acta Crystallogr.*, **A35**, 901–911.
Yanofsky,C., Horn,V., Bonner,M. and Stasiowski,S. (1971) *Genetics*, **69**, 409–433.

Received on July 6, 1989; accepted on October 19, 1989